Leaving no element of doubt: analysis of proteins using microPIXE

Elspeth Garman

Address: Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK.

E-mail: elspeth@biop.ox.ac.uk

Structure December 1999, 7:R291–R299

0969-2126/99/$ – see front matter

Introduction

The identification of atoms, anions and cations in protein structure electron density is often inferred from the temperature factor of the atom and from its coordination distances to surrounding amino acids. In many cases unambiguous identification is not possible without the use of additional techniques. Mass spectrometry analysis methods [1] have advanced significantly over the past few years and precise total mass measurements are now achievable. There is often ambiguity in the identification of bound atoms, however, because extra mass can arise from a combination of sources; for example, from covalent modification such as that of cysteine residues or glycosylation, and from partially occupied metal atom sites as opposed to totally occupied lighter atom sites.

One unique signature of an element that can be used in unambiguous identification is the energy of an X-ray emitted when an atomic electron undergoes an energy transition between its shell and a vacant electron site in a lower energy shell (e.g. for an M to L shell transition, sulphur gives a 2.3 keV X-ray). The energy resolution of modern single crystal solid-state X-ray detectors allows assignment of characteristic X-ray peaks in an energy spectrum to the various elements of the periodic table.

Emission of the characteristic X-rays from a sample can be induced by an incident beam of higher energy X-rays (X-ray fluorescence, XRF), electrons (electron probe X-ray microanalysis, EPMA) or by protons (proton-induced X-ray emission, PIXE). All three techniques have been used on a wide variety of biological specimens [2]. None of them has a clear cut advantage for all experimental cases, and in fact the various advantages and disadvantages make them complementary rather than competitive techniques.

Synchrotrons produce beams intense enough to make a microscopic probe for XRF, and the natural horizontal polarisation of the X-ray beam results in increased signal-to-noise ratios over non-polarised X-ray beams. The photoionisation cross-sections for K and L series X-rays in XRF are very low for the lightest elements (below atomic number $Z = 10$), but rise rapidly with increasing atomic number. In contrast, for ionisation of atoms with electrons (EPMA) and protons (PIXE), the cross-sections for interactions which subsequently cause X-ray production are high for elements with low atomic numbers but decrease with increasing atomic number. Thus for identifying heavier elements in samples from their X-ray fluorescence, the minimum detectable limit when using XRF is smaller than if using PIXE, whereas for light elements such as sodium and magnesium, PIXE has a lower detectable limit. Proton ion sources are more widely available than synchrotrons and proton microprobe beams can be focussed down to a smaller (below 1 μm) diameter than synchrotron-produced X-ray beams (minimum of ~20 μm), giving PIXE the better spatial resolution. Where spatial information on the distribution of elements in the sample is obtained by PIXE, the technique is known as microPIXE.

For electron microprobes, the 'Bremsstrahlung' background from the primary electron beam is much higher than for a proton beam, and thus the minimum detectable limit for EPMA is higher than for PIXE. Electrons of the energies used in electron microprobes penetrate ~2–3 μm into the sample, and there has been a long standing debate as to the exact shape of the 'tear-drop' volume of the sample in which X-ray emission is excited (e.g. [3]). For this reason accurate correction for self-absorption of X-rays in the sample is problematic. A proton microprobe beam penetrates much deeper into the sample (e.g. ~60 μm for a 3 MeV proton beam) and a linear integration of X-ray production along its path length is possible and more accurate. For EPMA studies, samples must be coated in a conducting layer of gold or carbon, whereas sample preparation for PIXE is more straightforward.

This paper reports a method for using a scanning proton microprobe (SPM) for PIXE analysis of protein samples. Using this technique, all elements apart from those lighter than sodium ($Z = 11$) can be detected at a minimum limit of one to ten parts per million by weight. PIXE allows the detection of many elements at once, and for proteins, following corrections, simple ratios can be taken to quantitate the analysis, obviating the need for absolute measurements (see later). If the protein sequence is known, information on the number of atoms per protein molecule can be obtained by using the sulphur concentration from the methionines and cysteines in the protein as an internal calibration. The wide ranging applications of this technique include the qualitative and quantitative identification of atoms/ions in reaction centres, active sites, metal-binding proteins and protein-bound DNA or RNA, the identification of putative

atoms and ions in the electron density of solved structures, and the determination of the degree of selenomethionine substitution in a protein.

Measurements are described on one virus and a range of proteins. They include an epidermal growth factor (EGF)-like domain of factor IX, a viral and a bacterial neuraminidase, protein phosphatase 1, a cell-cycle control protein and human immunodeficiency virus 1 (HIV-1) reverse transcriptase co-crystallised with an oligonucleotide. The results of some of these measurements are already in the literature [4–8], but the methodology of the technique has not been previously described. The protocol that has been developed is applicable to both dried protein solutions and crystalline protein samples.

There are at least 50 SPMs located around the world that could potentially be used to perform these analyses, and more are coming on line every year. The technique is thus of interest as it adds another option to the tools available to the structural biologist, providing a method complementary to the use of XRF and electron microprobes.

Ion beam analysis

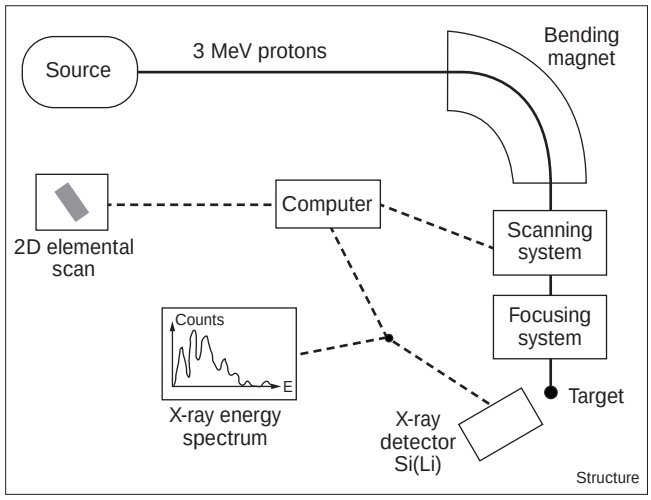
The scanning proton microprobe

An SPM has been developed at Oxford University [9]. Over the past ten years it has been used for the elemental mapping and analysis of a huge variety of samples from disparate fields of research. Elements in the sample can be identified at a minimum detectable limit of one to ten parts per million (ppm) by weight. The proton beam diameter achieved is routinely ~1 μm, making the SPM a true 'microprobe'.

The SPM is shown diagramatically in Figure 1. A 2–3 MeV proton beam from a small van de Graaff accelerator is focussed to a diameter of ~1 μm using high precision magnetic lenses (Oxford Microbeams, Oxford UK). The induced X-rays from elements in the target sample are detected using a lithium drifted silicon (Si(Li)) detector, which gives a voltage pulse proportional to the X-ray energy. Elements with a low atomic number give lower energy X-rays than those of higher atomic number; elements with an atomic number of less than 11 (sodium) produce X-rays that are absorbed in the front detector window (Ex < 0.7 keV) and are thus not detected. For the detection of light elements in the sample (e.g. sodium and magnesium) the energy of the proton beam can be lowered from 3 to 2 MeV. This reduces the maximum energy of the secondary electrons produced in the sample and thus reduces the Bremsstrahlung background.

A high precision computer-controlled scanning system allows the incident proton beam to be moved in x and y, so that X-ray emission can be induced at different positions on the sample.

Figure 1



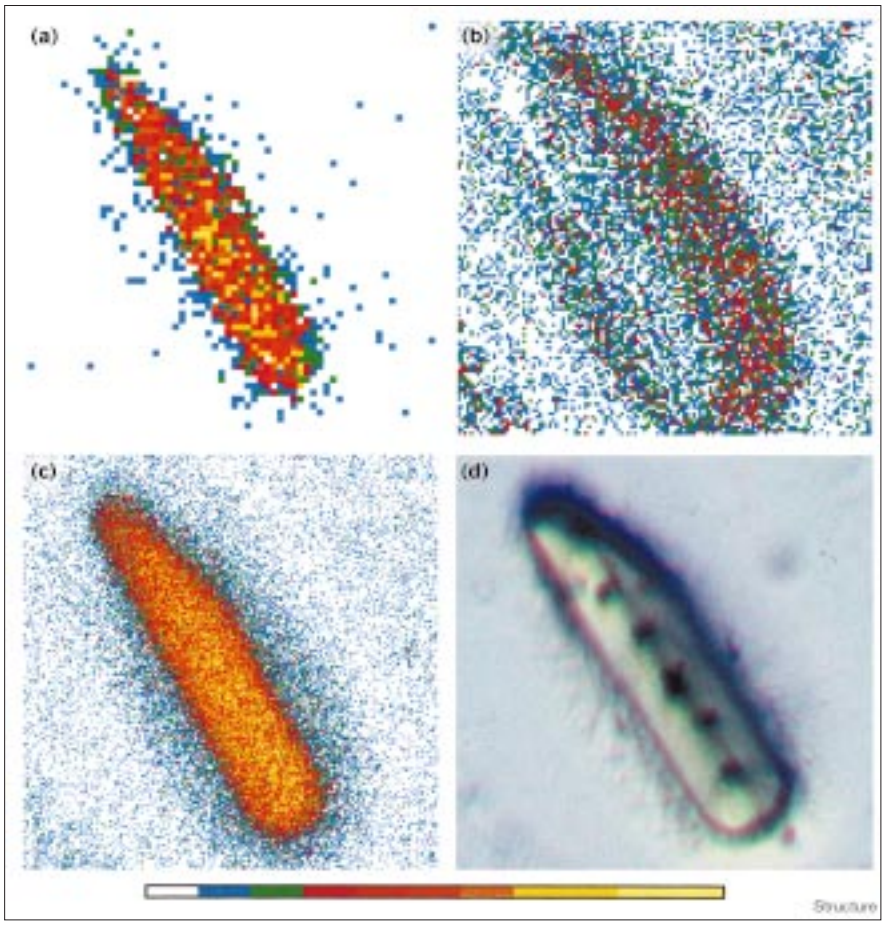
Scanning proton microprobe for PIXE analysis. Schematic of the experimental arrangement. Protons (2–3 MeV) emerge from the van de Graaff accelerator and are bent through 90° before being focused by high precision magnets onto the sample. The whole beamline is kept under high vacuum.

Elemental analysis with the SPM

The SPM can operate in two different modes, both of which were used in the measurements described here. The first of these is the two-dimensional (2D) sample mapping mode, where the proton beam is scanned in x and y over the sample and software windows are set on X-ray peaks of interest to build up 2D elemental maps of the sample. The area of the scan can be chosen according to the size of the sample. Figures 2 and 3 give examples of such maps obtained for protein crystals. This analysis is semiquantitative, in that from the total X-ray energy spectrum collected it can identify which elements are present and at roughly what level. For accurate quantitation, however, the SPM is operated in the second mode — probe analysis. In this mode spectra are collected for ~10 min each at selected (x,y) points both on and off the sample to obtain the background. These spectra are then analysed to give the concentration of the elements at that point. A 2D map is collected first and is then used for choosing appropriate (x,y) positions on the sample. The small physical dimensions of the beam enables the experimenter to achieve high (x,y) spatial resolution of the distribution of the various elements in the sample.

Once a spectrum from a point on the sample has been collected, in common with all other X-ray emission techniques, corrections have to be applied. For PIXE these corrections are more easily parameterised than for EPMA, as the volume of sample penetrated by the beam is known more precisely. The PIXE spectra are corrected and analysed using the software program GUPIX [10]. This

Figure 2



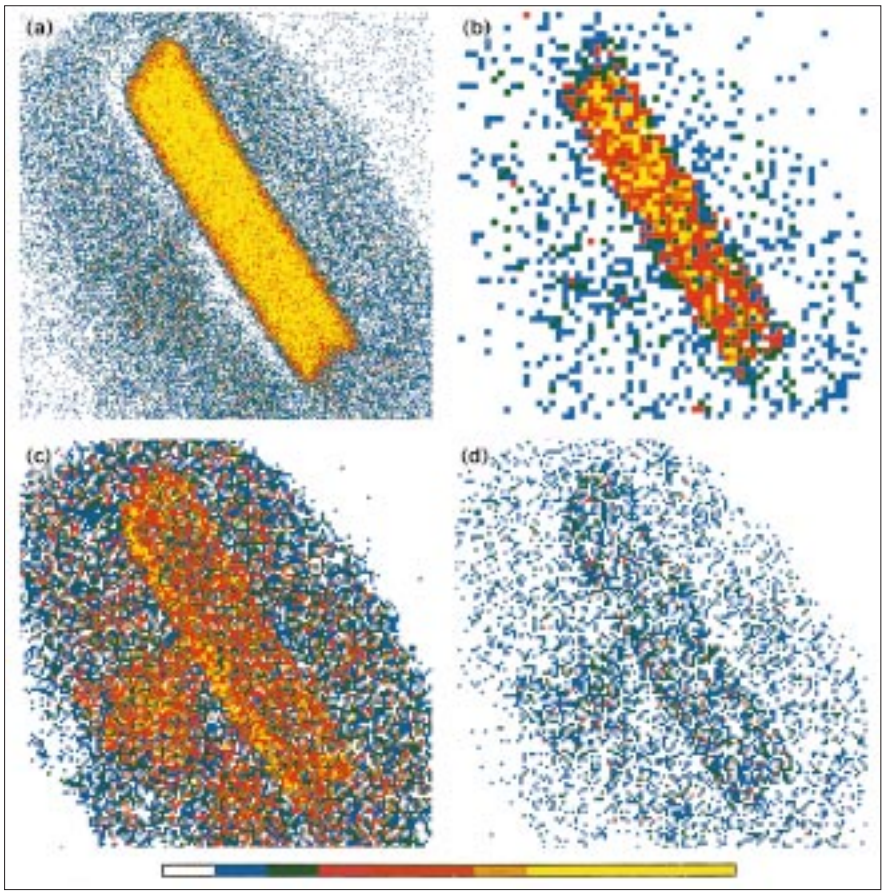
Example of a 200 × 200 μm two-dimensional PIXE scan giving elemental maps of a crystal of p13(suc1), a protein involved in cell-cycle control [13]. The crystal is surrounded by dehydrated mother liquor containing sodium cacodylate. The figure shows the colour-coded intensity range of the emitted X-rays for a particular energy (wavelength), corresponding to a unique element, over the grid which encompasses the crystal. The colours range from blue to yellow, representing low to high intensity emission, respectively. Elemental plots are shown for (a) zinc, (b) sodium, and (c) sulphur. In the zinc and sulphur maps the outline of the crystal is clearly visible, showing these elements to be present in the protein sample and not present in the surrounding area. The sodium map reflects the distribution of mother liquor (containing sodium cacodylate) around the crystal. (d) Light photo of the same crystal after the PIXE analysis. The five dark holes in the crystal are places where point spectra were collected; the proton beam has clearly bored through the crystal, and the resulting holes are much larger (about 8 μm) than the 1 μm beam diameter due to the spread of radiation damage.

allows the extraction of background-subtracted peak-fitted intensities which have been corrected for a number of factors: the change of ionisation cross-section as proton energy is lost in the sample, the change in X-ray self-absorption with proton penetration depth, the fluorescence of lower mass atoms induced by PIXE from higher mass atoms (secondary fluorescence), the detector response and the beam–sample–detector geometry. The correction procedure also accounts for spurious detector responses apparent in the spectrum, such as escape peaks and pile-up peaks.

It is possible to make these corrections in a reproducible and accurate way because the four main parameters that affect the overall X-ray yield all vary slowly with atomic number and are well defined. These four parameters are proton beam loss, the X-ray absorption in the sample as the X-rays come out, the ionisation cross-section as a function of energy, shell and atomic number, and the fluorescent yield. Following careful measurements by researchers over many years, the physics database now holds the necessary information for GUPIX to perform realistic corrections for quantitative PIXE analysis. This obviates the need for continual calibration with reference samples. The reliability of the correction procedure depends on a knowledge of the sample target thickness, especially for a thick target (> 60 μm) when the proton beam will not penetrate right through the sample. The target thickness can be obtained by detecting the Rutherford back-scattered protons and fitting the resulting spectrum (discussed later).

The correction procedure introduces systematic errors into the analysis which limit the eventual accuracy of the results [11]. These include inaccuracies in the assumptions that have to be made about the target flatness and homogeneity, numerical errors in processing the spectrum, limitations in the model used for the detector response function, and inaccuracies in the experimental values taken from the physics database (which are probably small compared with the other sources of error). These systematic errors combine with the error contribution from the counting statistics to give the final uncertainties in the extracted intensities of the X-ray peaks.

Figure 3



Example of a 500 × 500 μm two-dimensional PIXE scan showing elemental maps of a protein phosphatase 1 tungstate-bound crystal. Four elemental plots are shown (a) sulphur, (b) manganese, (c) tungsten, and (d) zinc. In the sulphur and manganese maps the outline of the crystal is clearly visible, showing these elements present in the protein sample and not present in the surrounding area. No significant levels of zinc are present. (Results taken from [7]).

The X-rays detected during PIXE analysis belong to the K, L and M series. For all elements lighter than zirconium, the lines from the K series are used for quantitative analysis. Above zirconium, the K line X-rays are rather high in energy, making the L series lines more appropriate for use in PIXE analysis.

A very important feature of PIXE analysis is that the measured X-ray signal does not depend on the chemical environment of the atoms.

Experimental method for protein analysis

Sample preparation

Sample holders are prepared beforehand by attaching a taut thin mylar film (~1 or 2 μm thick) over a 0.5 cm diameter hole in a 2.5 × 1.25 cm aluminium frame with glue (liquid cement for plastics). Samples of proteins in solution (~0.5 μl) are pipetted onto the film; for crystals, a cryo-loop can be used for the transfer. The samples are covered to prevent contamination and left to dry out naturally overnight. If crystals have been grown in mother liquor containing either any of the elements of interest or elements with similar atomic number, thorough and careful washing in de-ionised water prior to mounting is required. For example, the buffer MES contains sulphur which gives a background under the sulphur peak originating from cysteine and methionine residues in a protein, thus leading to the wrong relative ratio of sulphur and the atom of interest. Likewise, any sodium chloride in the sample gives a chlorine peak in the X-ray spectrum, which if strong will interfere with an accurate extraction of the intensity in the neighbouring sulphur peak. For all crystalline samples, unwashed and washed crystals are now routinely analysed so that the results can be compared.

In the SPM, four sample holders at a time are attached to a metal ladder and pumped down in the target chamber of the beamline to a vacuum of ~10^{-6} τ, which takes around 10 min.

Elemental analysis of protein samples

The sample is aligned in the beam using a camera with a view into the evacuated target chamber. A coarse 2D scan (e.g. 1 × 1mm) is collected, and after location of the sample a finer scan is performed (e.g. 200 × 200 μm; see Figures 2,3). Four or five different (x,y) positions on the

sample, one or two on the mother liquor (if the sample is a crystal) and one on the mylar film are then selected for the collection of point spectra for quantitative analysis. Adequate counting statistics can usually be obtained by collecting data at each point for between 10 and 15 min.

Knowing the sequence of the protein, the number of cysteine and methionine residues (which each contain one sulphur atom) can be used to normalise the amount of sulphur detected in the PIXE measurement, and thus to calculate the number of atoms of interest per protein molecule. The sulphur content of the protein gives a very convenient 'internal calibration' which does not depend on knowledge of the target thickness, total beam charge, detection solid angle, sample thickness and sample composition; all these would be required for an absolute measurement. For DNA or RNA samples, if the number of bases is known, the phosphorus concentration can be used to obtain the internal calibration.

The concentrations of all the selected elements of interest, in parts per million, are extracted from the X-ray spectrum taken at a particular (x,y) position by the peak-fitting software GUPIX [10]. This extracted concentration is the true concentration multiplied by a normalising factor. Thus the ratios of concentrations of the element u of interest ($C(u)$) and of sulphur ($C(s)$) (or phosphorus for a DNA or RNA sample), can be obtained without knowing the normalising factor. This is a crucial feature of the technique in allowing accurate quantitative measurements for protein samples.

The concentrations, ($C(u)$) and ($C(s)$), can then be converted into the number of atoms of u per protein molecule ($N(u)$) by dividing by the mass ratio of the element u and sulphur ($M(u)/M(s)$), and multiplying by the number of sulphur atoms per protein molecule ($N(s)$) (or number of phosphorus atoms in the case of DNA or RNA).

$$N(u) = [C(u)/C(s)] \times [M(s)/M(u)] \times N(s)$$

This calculation is repeated for each point spectrum measured on the sample, and for the background (mylar film and, in the case of crystals, the mother liquor dried round the crystals). An average of the 'on-sample' results is then taken and a variance calculated.

Calibration

The SPM has undergone extensive calibration tests to validate the methodology, software, and results from many different sample types. For protein crystals, calibration was carried out with $CaSO_4$ crystals and also with $CaSO_4$ powder. $CaSO_4$ was chosen because the sulphur and calcium were a reasonable mimic of a real protein measurement where the number of calciums per protein molecule were being investigated. The above experimental method was used, and the relative concentrations of calcium and sulphur were measured at four points for three different sized crystals. A ratio of 1:1 (± 6%) was obtained. Calcium and sulphur are close in atomic number, so systematic errors imposed by the integration and correction software are minimised. Thus the error of ± 6% probably represents a minimum achievable error in these measurements.

Results

The results of PIXE measurements performed using the Oxford SPM over the past seven years are summarised in Table 1. For crystal samples, measurements were made on washed and unwashed crystals and then compared: those for washed samples are presented (see later discussion). The elemental analysis of each sample was prompted by a variety of different problems, ranging from checking crystals for the incorporation of co-crystallised components, to identifying anions and cations in the electron density of solved structures. Thus in order to illustrate some potential applications of the microPIXE technique, the reason for some of the analyses and the implications of the results for the particular project will be described briefly.

HIV-1 reverse transcriptase

In an effort to obtain diffraction-quality crystals of the HIV-1 reverse transcriptase heterodimer (66 kDa/51 kDa), co-crystallisations with different length oligonucleotides were attempted. The crystals resulting from the addition of a DNA 22:9-mer oligodeoxynucleotide were analysed by microPIXE to check (from the phosphorus signal) that the oligonucleotide had been incorporated and to quantify how many molecules per heterodimer were present (from the phosphorus to sulphur ratio). The results indicated that 1.2 (± 0.5) oligonucleotides were present per heterodimer, which was consistent with each heterodimer possessing one competent active site [5]. This conclusion has since been borne out by the results of the structure determination. In this particular PIXE measurement the error was somewhat higher than usual, because sulphur (Z = 16) and phosphorus (Z = 15) have neighbouring atomic numbers and the tails of the X-ray peaks in the PIXE spectrum therefore overlap one another, giving a larger uncertainty in their individual intensities than if comparing elements further apart in atomic number.

Viral neuraminidase

In viral neuraminidase, calcium is known to increase the initial rate of activity of the enzyme, and putative calcium ions have been observed in the X-ray structures where the ion is thought to have a stabilising effect on the active site but does not have a direct catalytic role. The presence of calcium was unambiguously confirmed by microPIXE analysis of influenzae subtype N8 neuraminidase crystals [4]. The 80 kDa bacterial neuraminidase from *Vibrio cholerae* requires calcium for activity, and again calcium was found in the crystals by microPIXE analysis; calcium was later

Table 1

The results of PIXE analyses performed using the Oxford SPM.

Sample	Elements of interest	Result atoms/molecule	Reference
Reverse transcriptase + DNA oligodeoxynucleotide 22:9mer	P	1.2 (± 0.5) oligonucleotide per dimer	[5]*
Viral neuraminidase (subtype N8)	Ca	Present	[4]*
Bacterial neuraminidase (VC)	Ca	Present	[12]
Glucose dehydrogenase	Zn	Present	[15]
Epidermal growth factor like domain of factor IX	Ca	1.7 (± 0.2)	[6]*
p13(suc1) cell-cycle control protein	Zn	1.3 (± 0.2)	[13]
Protein phosphatase 1 (native)	Mn	0.93 (± 0.08)	[7]*
	Fe	0.46 (± 0.05)	
Protein phosphatase 1 (with bound tungstate)	Mn	0.92 (± 0.08)	[7]*
	Fe	0.45 (± 0.05)	
Reduced desulphoferrodoxin	Fe	1.1 (± 0.1)	(A Coelho, unpublished data)
	Ca	0.5 (± 0.05)	
Oxidised desulphoferrodoxin	Fe	2.3 (± 0.15)	
	Ca	1.35 (± 0.15)	
Class I ribonucleotide reductase R2 subunit mutant D84A	Fe	bdl†	
	Zn	0.23 (± 0.03)	
	Hg	1.5 (± 0.12)	[16]
Class III anaerobic ribonucleotide reductase	Fe	bdl†	
	Zn	1.5 (± 0.3)	
	Cu	3.0 (± 0.5)	[17]
Blue tongue virus (BTV) core	Counter ions	Ca, Zn at low concentration	[8]*
VP7 capsid protein from BTV	Zn	Present	[18]
CD80 (B7)	Se	S:Se = 1.8 (± 0.2)	(S Davis, S Ikemizu, DI Stuart and EY Jones, unpublished data)
Tyrosinase	Cu	Present	(M Wormald, unpublished data)

*SPM results reported in cited reference. †bdl, the concentration was below the minimum detectable limit.

identified in the electron density of the crystals when the structure was solved [12]. In both cases no other metal ions were detected at significant levels in the PIXE analysis.

Factor IX EGF-like domain

Crystals of the EGF-like domain of factor IX were analysed by microPIXE and found to have 1.7 (± 0.2) calcium atoms per protein monomer. When the structure was solved [6], the electron-density maps showed that the molecule associated in dimers and that there was one calcium atom per monomer and one calcium atom at the interface of each dimer, giving an average of 1.5 calciums per monomer in concurrence with the PIXE results.

p13(suc1)

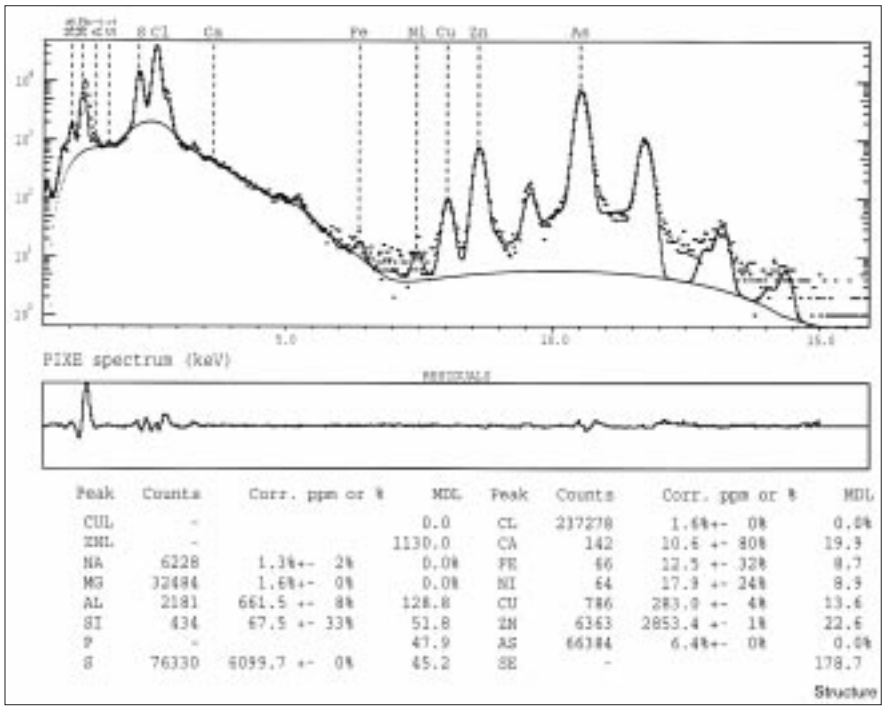
The electron density of p13(suc1), a protein involved in cell-cycle control, showed three zinc ions bound per suc1 dimer: two at each monomer–monomer interface and one at a crystal-lattice contact (i.e. an average of 1.5 zinc ions/monomer) [13]. Initial microPIXE analysis gave inconsistent results that were caused by the mother liquor, which was buffered with Tris-HCl. The chlorine peak from this interfered with an accurate measurement of the sulphur concentration. When the crystals were transferred to sodium cacodylate buffer, in which they were known to be stable, more consistent values were obtained, although the arsenic from the cacodylate gave some background near the zinc peaks. The crystals were then thoroughly washed in de-ionised water and subjected to microPIXE analysis, which gave 1.3 (± 0.2) zinc ions/monomer, agreeing with the structural results. Figure 2 shows the elemental maps for sulphur, zinc and sodium obtained for the crystal in cacodylate. The figure also shows a conventional photograph of the same crystal after microPIXE analysis, showing holes at the points where the proton beam was incident for collection of the five-point spectra. Figure 4 shows an X-ray spectrum, plotted on a log scale, from one of the five points collected on the crystal with the GUPIX fit superimposed: the arsenic and residual chlorine peaks are seen to be strong.

Human phosphatase 1

The structure of the catalytic subunit of human phosphatase 1 (PP1) was solved by a multiwavelength anomalous diffraction (MAD) experiment on a tungstate-bound crystal [7]. PP1 is a serine/threonine protein phosphatase and has a dinuclear ion centre situated at the catalytic site which binds the phosphate moiety of the substrate. microPIXE was used to identify the ions bound in the centre as iron and manganese, and consistent measurements were obtained from a native crystal and one with bound tungstate. Half an iron and one manganese ion were identified for each PP1 molecule. The electron-density height at one metal site was half that of the other

Figure 4

X-ray spectrum taken from a point on the same p13(suc1) crystal illustrated in Figure 2. The dots represent the experimental data (X-ray counts in the Si(Li) detector) and the solid line represents the GUPIX fit. The software extracts the intensities by peak-fitting and integration of the counts in the peaks to obtain, following correction (see text), the weight in parts per million (ppm) or as a percentage. Note the log scale and the large arsenic peak from the cacodylate buffer.



Peak	Counts	Corr. ppm or %	MDL	Peak	Counts	Corr. ppm or %	MDL
CUL	-		0.0	CL	237278	1.6%+- 0%	0.0%
ZNL	-		1130.0	CA	142	10.6 +- 80%	19.9
NA	6228	1.3%+- 2%	0.0%	FE	66	12.5 +- 32%	8.7
MG	32484	1.6%+- 0%	0.0%	NI	64	17.9 +- 24%	8.9
AL	2181	661.5 +- 8%	128.8	CU	786	283.0 +- 4%	13.6
SI	434	67.5 +- 33%	51.8	ZN	6363	2853.4 +- 1%	22.6
P	-		47.9	AS	66384	6.4%+- 0%	0.0%
S	76330	6099.7 +- 0%	45.2	SE	-		178.7

site, indicating that the occupancy of one was half that of the other if the two ions were of similar mass. This matched the relative ratios given by the PIXE analysis, allowing tentative assignment of the sites to the specific ions. Another alternative is that both sites contain a mixture of both ions, and there is currently no method to distinguish between the two possibilities.

Blue tongue virus core

Washed, inactivated crystals of blue tongue virus (BTV) cores were subjected to microPIXE analysis in an effort to identify any counter ions that could neutralise the negatively charged double-stranded (ds)RNA inside the core; dsRNA has been modelled into the electron density of the X-ray structure [8]. In the PIXE analysis, insufficient metal ions were present to act in any substantial way to neutralise the dsRNA: calcium and zinc were present at only one-hundredth of the level of phosphorus. This analysis suggests that an organic cation might be present [8].

Selenomethionine substitution

Selenomethionine-substituted CD80(B7) protein was produced by mammalian expression and there was concern as to the level of labelling (S Davis, S Ikemizu, DI Stuart and EY Jones, personal communication). Glycosylation of the protein rendered mass spectrometry less suitable than PIXE for checking the selenium content prior to MAD data collection. The ratio of cysteine to methionine residues in the protein was known to be 4:6. When the sulphur to selenium ratio of the crystals was measured using microPIXE, a ratio of 1.8 (± 0.2) sulphur atoms for every selenium atom was obtained, indicating that 60% (± 5) of the methionine residues in the protein were labelled with selenium and thus MAD data collection was a possibility.

Discussion

The results detailed above clearly indicate that microPIXE analysis of proteins in solution and crystalline protein samples can provide unambiguous answers to specific questions. The elemental information, especially the quantitation at an accuracy of 6–10%, is currently not easy to obtain by other techniques. The 1 μm diameter beam allows clear resolution of the sample on the mylar film and reliable analysis of small samples (e.g. crystals with a maximum dimension of 20 μm).

For protein solution samples, a major advantage of microPIXE over other analytical techniques is that the protein concentration does not have to be known, as the internal calibration provided by the cysteine and methionine residues makes this unnecessary. Protein samples can be analysed before crystallisation for the purpose of general characterisation, and to check for the presence of expected/unexpected metal atoms, or, for instance, to investigate the changed metal-binding properties of a mutant compared to native protein.

Two main problems were encountered in characterising the microPIXE technique for use on protein crystals. The

first was that the mother liquor (e.g. calcium chloride) often contains elements that interfere with a clean X-ray spectrum, because the intense peaks from elements such as chlorine and calcium produce spurious responses in the Si(Li) detector. In particular, they result in the electronic pile-up of pulses, which then give peaks at high X-ray energy, and in escape peaks being detected at lower energy. These extra peaks appear in the spectrum, sometimes overlapping the peaks of interest and making them impossible to integrate accurately. It was thus found to be vital to wash the crystals prior to microPIXE analysis, and to compare the microPIXE results from washed and unwashed crystals. Although a spectrum is always collected on the dried mother liquor around the crystal, it is not possible to use this to correct accurately for the mother liquor present in the crystal. For such correction, knowledge of the solvent content and composition, and of the thickness of the dried liquid layer on the surface of the crystal would be required (i.e. there is no internal calibration that can be applied to the mother liquor contribution). The second problem encountered was that crystals occasionally crack as they dry onto the mylar film, and if the proton beam hits a crack spurious results are obtained. This is due to lack of precise knowledge of the self-absorption of X-rays emanating from the walls of the crack. This self-absorption is more significant for the lighter elements (e.g. sodium and chlorine) as they emit lower energy X-rays. The software program GUPIX models this absorption, but cannot do so adequately for very uneven samples. Thus after inspection of the (x,y) maps, at least four flat and uniform places on the sample were chosen for point spectrum analysis. The results from these points could then be checked for consistency.

The self-absorption of X-rays in the sample is modelled by GUPIX using information input by the experimenter on the approximate sample thickness and composition. In order to measure the sample thickness, the Oxford SPM has the capability of detecting the Rutherford backscattered (RBS) protons in a solid-state silicon detector placed inside the evacuated target chamber. The proton energy spectrum from scattering caused by light elements in the sample (e.g. carbon, oxygen and nitrogen) can then be fitted to obtain the target thickness and thus a better correction for the X-ray self-absorption [14]. For proteins, which contain mainly these light elements, use of RBS for determining the target thickness has proved an accurate method. It is now used routinely, and hence PIXE analysis gives more consistent and interpretable results. Accurate knowledge of the target thickness is far more important for a proton beam probe than for an electron beam probe, because the range of the protons is much higher and is of the order of the target thickness. Thus X-rays are produced all along the beam path and self-absorption must be considered correctly to obtain accurate results.

Unexpected elements are often found to be present in the samples. For instance, if a nickel column is used in the protein purification nickel is sometimes evident. A knowledge of the history of the protein is thus an advantage when interpreting the data.

If the protein contains no sulphur, because the sequence lacks cysteine and methionine residues, accurate quantitation in terms of the number atoms of interest present per protein molecule is extremely difficult to extract. Any quantitation then depends on knowledge of absolute values of the protein concentration (for liquid samples), beam current, crystal composition and thickness (for crystals), and so on. For protein samples this has not yet been attempted, as it is only necessary in the absence of any sulphur content and none of the samples studied thus far has fallen into this category.

The microPIXE analysis of elements bound to proteins cannot distinguish between one fully occupied ion site and two partially occupied ion sites, as it only measures the total concentration present per protein molecule.

Conclusions

microPIXE analysis using an SPM on solution residue and crystalline protein samples is capable of the unambiguous identification and quantification of anions/cations/atoms. The number of anions/cations/atoms per protein (or DNA/RNA) molecule can be determined to an accuracy of 6–10% providing the sequence is known and the protein contains at least one cysteine or methionine residue. Often the anions/cations/atoms in proteins have never been positively identified; previously circumstantial evidence has been relied upon to give putative identification. The analysis can also be used to check the degree of selenomethionine substitution prior to attempting MAD experiments, and to establish DNA/RNA binding to proteins before structural analysis. The need to wash crystal samples thoroughly before microPIXE analysis usually renders the technique unsuitable for checking heavy-atom binding prior to multiple isomorphous replacement (MIR) experiments.

Although the availability of state of the art proton microprobes is limited, the simple sample preparation procedure enables samples to be sent to an analytical service (e.g. Spectrum Microanalysis Ltd, Oxford, UK). The presence of known amounts of sulphur in proteins provides fortuitous internal calibration for PIXE, establishing the technique as a very useful addition to the analytical tools available to the structural biologist.

Acknowledgements

I am deeply grateful to Geoff Grime, who has been involved in the development and operation of the Oxford SPM ever since its inception. He has given continual support and advice on the application of the technique to protein samples, and provided constructive criticism of this article. I also thank Frank Watt for early encouragement, and various members of the

Laboratory of Molecular Biophysics in Oxford who have helped the development of the technique by enthusiastically allowing measurements on their samples. Thomas Schneider kindly commented on the manuscript and Stephen Lee assisted with the figures. EG is supported by the UK Medical Research Council.

References

1. Chait, B.T. (1994). Mass spectrometry – a useful tool for the protein X-ray crystallographer and NMR spectroscopist. *Structure* **2**, 465-467.
2. Llabador, Y. & Moretto, P. (1996). *Applications of Nuclear Microprobes in the Life Sciences.* World Scientific, Singapore.
3. Merlet, C. (1992). Evaluation and description of a new correction procedure for quantitative electron probe microanalysis. *Inst. Phys. Conf. Ser.* **130**, 123-126.
4. Taylor, G., Garman, E., Webster, R., Saito, T. & Laver, G. (1993). Crystallization and preliminary X-ray studies of influenza subtypes N5, N6, N8 and N9. *J. Mol. Biol.* **230**, 345-348.
5. Jones, E.Y., *et al.,* & Stammers, D.K. (1993). The growth and characterization of crystals of immunodeficiency virus (HIV) reverse transcriptase. *J. Crystal Growth* **126**, 261-269.
6. Rao, Z., Handford, P., Knott, V., Mayhew, M., Brownlee, G.G. & Stuart, D. (1995). Crystallization of a calcium-binding EGF-like domain. *Acta Crystallogr. D* **51**, 402-403.
7. Egloff, M.-P., Cohen, P.W.T., Reinemer, P. & Barford, D. (1995). Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *J. Mol. Biol.* **254**, 942-959.
8. Gouet, P., *et al.,* & Mertens, P.P.C. (1999). The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. *Cell* **97**, 481-490.
9. Grime, G.W., Dawson, M., Marsh, M., McArthur, I.C. & Watt, F. (1991). The Oxford submicron nuclear microscopy facility. *Nucl. Instr. Methods Phys. Res. B* **54**, 52-63.
10. Maxwell, J.A., Teesdale, W.J. & Campbell, J.L. (1995). The Guelph-PIXE software package-II. *Nucl. Instr. Methods B* **95**, 407.
11. Grime, G.W. (1999). High energy ion beam analysis methods. In *Encyclopedia of Spectroscopic Techniques.* Academic Press, in press.
12. Crennell, S., Garman, E., Laver, W.G., Vimr, E. & Taylor, G.L. (1994). Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. *Structure* **2**, 535-544.
13. Endicott, J.A., *et al.,* & Johnson, L.N. (1995). The crystal structure of P13 (suc1), a p34(cdc2)-interacting cell cycle control protein. *EMBO J.* **14**, 1004-1014.
14. Grime, G.W. (1996). The 'Q factor' method: quantitative micoPIXE analysis using RBS normalisation. *Nucl. Inst. Methods Phys. Res. B* **109/110**, 170-174.
15. John, J., Crennell, S.J., Hough, D.W., Danson, M.J. & Taylor, G.L (1994). The crystal structure of glucose dehydrogenase from *Thermoplasma acidophilum*. *Structure* **2**, 385-393.
16. Nordlund, P., Sjöberg, B.-M. & Eklund, H. (1990). Three-dimensional structure of the free radical protein of ribonucleotide reductase. *Nature* **345**, 593-598.
17. Logan, D.T., Andersson, J., Sjöberg, B.-M. & Nordlund, P. (1999). A glycyl radical site in the crystal structure of a class III ribonucleotide reductase. *Science* **283**, 1499-1504.
18. Basak, A.K., Grimes, J.M., Gouet, P., Roy, P. & Stuart, D.I. (1997). Structures of orbivirus VP7: implications for the role of this protein in the viral life cycle. *Structure* **5**, 871-883.